

PAPER

INTELLIGENT CLASSIFICATION OF LEARNER'S COGNITIVE DOMAIN USING BAYES NET, NAÏVE BAYES, AND J48...

Intelligent Classification of Learner's Cognitive Domain using Bayes Net, Naïve Bayes, and J48 Utilizing Bloom's Taxonomy-based Serious Game

<http://dx.doi.org/10.3991/ijet.v10i2.4451>
I Nyoman Sukajaya^{1,2}, I Ketut Eddy Purnama¹ and Mauridhi Hery Purnomo¹¹ Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia² Universitas Pendidikan Ganesha, Singaraja, Bali, Indonesia

Abstract—Lately, personalized learning approach attracted the attention of so many researchers due to its capacity to improve the quality of educational system. This approach provides opportunities to maximize the potential of all students based on their profile. This indicates the necessity of grouping learners' profiles appropriately in order to optimize contribution of personalized learning approach in achieving learning objective. The problems occurred when classifying learners' profile especially when dealing with large number of learners, restricted time to classify, and requirement of authentic data. To solve the problem, we proposed the implementation of Bloom's taxonomy-based serious game as an assessment tool replacing paper-based tool for the gameplay data collection. Three different methods namely: BN, NB, and J48 were implemented to obtain the highest accuracy of classification. Our study finds that the NB classifier gives the highest percentage accuracy that is 92.31%. This classifier has the similar accuracy with BN but with lower error rate. In view of the strength of agreement, the result is categorized Very Good ($\kappa = 0.85$).

Index Terms—Bloom's Taxonomy, Classification, Learner's Cognitive Domain, Serious Game.

I. INTRODUCTION

A well understanding of learners' characteristics will guide teachers in designing a good lesson plan since good lesson plan supports creation of joyful learning environment such that learning objective achieved effectively. Ideally, learners learn at their own characteristics. That is why the difference of learners' characteristics should be involved in designing a lesson plan.

Learners' characteristics were composed by a variety of aspects such as: personality, academic, social/emotional, and cognitive that is individualized [1]. Teachers were in charge of gathering authentically learners' information to profiling learners precisely. Besides, information must be gathered quickly. Delay in gathering learners' information has impact on delay of obtaining decision which is required in the improvement of lesson plan. This condition tends to widen bias of achievement of the learning objectives.

In practice, gathering learners' information is typically performed at assessment process. So, mistakes taking

place in the assessment activities have a major impact on the achievement of the learning objectives.

Although assessment process is an important part of learning process, it is not always simple to be performed [2]. Some of the conditions that constrain the assessment process are: large number of assessment's attributes involved, big size of class learners, and restricted time of assessment process. Furthermore, in case learners' information has been gathered, teachers are still made to classify a learner into an appropriate class.

The development of the game technology has attracted researchers to examine the opportunity of its application as an educational assessment tools [3, 4]. One such game among others believed as developing technology is a serious game [5]. One that distinguishing serious game of entertainment game is the existence of assessment component [6, 7] which is really required on education [3]. Serious game provides an opportunity for players to transferring technology and recording player behavior while playing the game [8]. The history of the application of game technologies in education can be tracked at [4, 9, 10, 11, 12, 13, 14].

Bloom's taxonomy theoretically explains that learners' mastery in a cognitive domain was really affected by their mastery on the previous domains [14]. It implies that learning process done gradually from the simplest to the most complex. Mechanism established in Bloom's taxonomy-based learning should be well adopted in setting the games' level of difficulty so as to avoid players' early frustration or boredom in playing the game.

Currently, there is no publication related to the utilization of Bloom's taxonomy-based serious game for the collection of data which is required in learners' classification. Therefore, this paper proposed the implementation of Bloom's taxonomy-based serious game assessment to replace the paper-based assessment for data collection. Data were gathered as information trail of Bloom's taxonomy-based serious game players. Learners playing the games were in charge of solving nine challenges distributed in the first three levels of cognitive domain of Bloom's taxonomy. Further, we conducted classification of players' cognitive domain by applying three methods namely: Bayes Net, Naïve Bayes, and Tree J48 to obtain the highest accuracy of players' classification.

The rest of this paper describes: literature review, Bloom's taxonomy-based serious game (BoTySeGa), research methodology, results and discussion, and concluding remarks.

II. LITERATURE REVIEW

A. Bloom's Taxonomy Model

In 1950 Benjamin Bloom developed a learning taxonomy which was then known as a Bloom's taxonomy. Taxonomy itself is defined as a classification device of learner's cognitive process which is involved in learning objective [14, 15]. Bloom classified learning process into three psychology domain i.e.: cognitive, affective, and psychomotor. Cognitive domain relates with information processing, affective domain deals with attitude and reasoning, and psychomotor relates with manipulation or physical skill [16].

Bloom's taxonomy views the cognitive skill of learners. He divides cognitive skill of learners in learning process into six different categories namely: knowledge (C1), comprehension (C2), application (C3), analysis (C4), synthesis (C5), and evaluation (C6). The categorization is ordered from the simple/low to complex/high thinking skill. Each category is stated using noun word (Fig. 1).

Further, several aspects, such as: cognitive ability, its brief description, synonyms, illustration of verbs used in assessment and level of difficulty of each of cognitive domain of Bloom's taxonomy are described at Table I.

B. Classification Methods

In this subsection we briefly describe three different methods of classification, namely: Bayes Net, Naïve Bayes, and Tree J48.

Bayes Net (BN), also known as Bayesian Networks, is a graphical model that is widely used to model industrial problems due to its probabilistic and graphical capabilities in reasoning under uncertainty. In education, BN can be used to represent the states of student ability.

BN relies on Bayes' rule [16]. The following is the definition of Bayes' rule:

- Given n attributes A_i which taken values a_i where $i=1, 2, 3, \dots, n$.
- Given C as class label and $E = (a_1, a_2, a_3, \dots, a_n)$ as unclassified test instance. E will be classified into class C with the maximum posterior probability. Bayes' rule for the classification is:

$$P(C | E) = \arg \max_C P(C)P(E | C) \quad (1)$$

Other Bayesian classifier which also known as the state-of-the-art Bayesian classifier is Naïve Bayes (NB). This classifier has been proven as one of the most computationally effective, efficient, and simple algorithm for machine learning and data mining application. NB classifier assumes that all attributes within the same class are independent given the class label. Based on this assumption, the Bayesian rule has been modified as follows to define the NB rule;

$$P(C | E) = \arg \max_C P(C) \prod_{i=1}^n P(A_i | C) \quad (2)$$

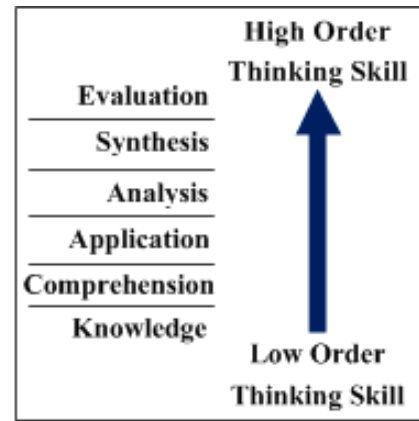


Figure 1. Bloom's Taxonomy (cited from [17])

TABLE I.
BLOOM'S TAXONOMY WITH APPROPRIATE SYNONYMS AND ILLUSTRATIVE VERBS

Objective	Definition	Synonyms	Illustrative verbs	Level
Evaluation	Judging the value of the system based on given criteria	Estimate Assessment	Justify; conclude; evaluation; verify; confirm; determine; analysis	Highest level dependent on students reasoning ability
Synthesis	Putting together elements/parts to form a system	Combination Fusion Creation	Generate; combine; construct; formulate; propose; assemble; design; predict; improve	High order question (HOq)
Analysis	Breakdown of a system into its elements/parts	Study Scrutiny Breakdown	Distinguish, compare; contrast; classify; categorize	HOq
Application	The use of abstractions in particular and concrete situations	Use Purpose Appliance	Change; demonstrate; modify; solve; use; show; calculate	HOq
Comprehension	Translation, interpretation and extrapolation of elements/parts	Understanding Grasp	Explain; convert; estimate; rearrange; summarize; derive; review; relate	Low order question (LOq)
Knowledge	Recall or recognition of specific elements/parts	Information Facts Data	Name; list; State; define; describe; label; sketch; discuss; identify; select; insert; complete	Lowest level dependent on students memory ability

J48 classification algorithm is the enhanced version of C4.5 decision tree and has been developed to generate a pruned or un-pruned C4.5 decision tree[18][19]. J48 builds decision trees from a set of labeled training data using the concept of information entropy. It uses the fact that each attribute of the data can be used to make a decision by splitting the data into smaller subsets. J48 examines the normalized information gain (difference in entropy) that results from choosing an attribute for splitting the

data. To make the decision, the attribute with the highest normalized information gain is used. Then, the algorithm recurs on the smaller subsets. The splitting procedure stops if all instances in a subset belong to the same class. Then a leaf node is created in the decision tree telling to choose that class. But it can also happen that none of the features give any information gain. In this case J48 creates a decision node higher up in the tree using the expected value of the class. J48 can handle both continuous and discrete attributes, training data with missing attribute values and attributes with differing costs. Further it provides an option for pruning trees after creation.

III. BLOOM'S TAXONOMY-BASED SERIOUS GAME

A. Game Scenario of BoTySeGa

The scenario of BoTySeGa adopts a story of a famous Indonesian folklore entitled Cupak Grantang. BoTySeGa is an adventure game telling the Grantang's adventure saving a princess from a giant named Menaru. A player acts as an agent who controls action of playing character to get successful in saving the adventure mission. The player can navigate the character in four different directions: left, right, up and down. Besides, he is also able to choose solutions of a challenge, and access Help before submitting a solution. The data recorded utilizing BoTySeGa are described in the section of Data Type and Value of Gameplay Data.

B. Design BoTySeGa

BoTySeGa is an assessment serious game which is developed to assess learners' mastery on mathematics subject matter of 5th grade elementary school learners. We involve parallelograms knowledge into BoTySeGa challenges. The Game is a single player game which is developed by author implementing a development software package RPGXP. Genre of BoTySeGa is adventure game and setting language used is Indonesian.

Player is in charge of playing character's tasks to solve the number of nine challenges distributed into the first three levels within the cognitive domain of Bloom's taxonomy. Only players who send the right solution on the third problem in each level may proceed to the next level of the game. Otherwise, the game is over. Those, who solve all problems rightly, are the winner. Challenges raised at BoTySeGa are associated with standard of competency which should be mastered by learners in the subject of parallelograms. The level of difficulty of the problems are categorized into three, namely: low (L), medium (M), or high (H). Determination of the level of problem's difficulty that is raised as a current challenge is based on the player's experience in the previous challenge (Fig. 2). Mapping of knowledge distribution in BoTySeGa is shown in Table II.

The players' experience which is recorded as BoTySeGa players' data are: score, completion time of challenge, and number of accessing Help. The other data recorded are players' achievement on the subject matter. The fourth data are recorded by way of using paper-based test. The players are directed to type his/her identity: name and registration number before playing BoTySeGa. This identity is used as a player's data key. BoTySeGa is also equipped with Help feature. Players can access this feature before they decide a solution for a problem in the

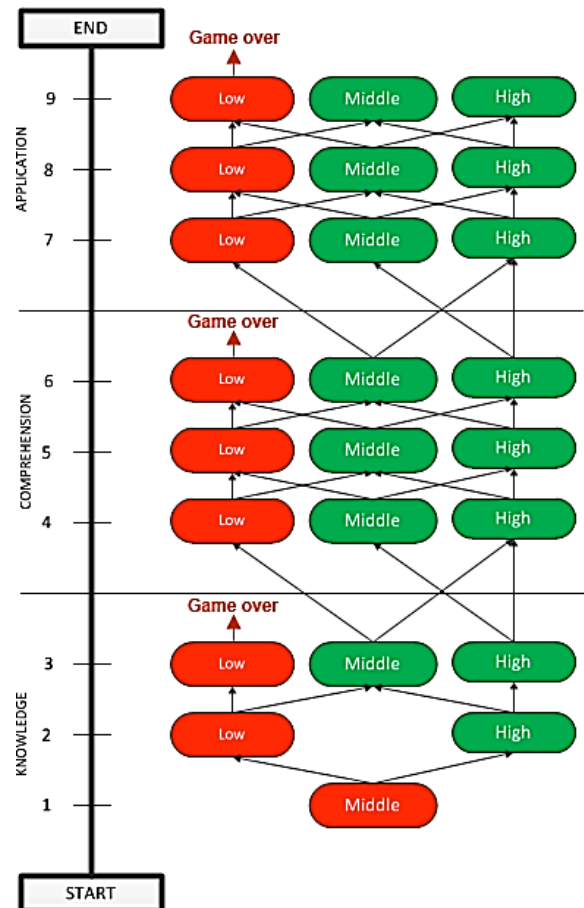


Figure 2. Rule of Auto leveling level of difficulty problem

TABLE II.
MAPPING OF KNOWLEDGE DISTRIBUTION IN BOTYSEGA

Probably Indicators	C1			C2			C3		
	L	M	H	L	M	H	L	M	H
Both pairs of opposite sides are parallel	1	17							
Both pairs of opposite sides are congruent	3	25	27	14, 19	7	8	35, 41		40
Both pairs of opposite angles are congruent	2	15	11	23, 24	16, 18	22	30	32	42
Consecutive angles are supplementary	6	24	31		9	13	37	39	43
One pair of opposite sides is congruent and parallel		5	20		29	33		38	44
Diagonal bisects each other	4		10	12		36	21	26, 28	45
Number of items	5	5	5	5	5	5	5	5	5

game. Fig. 3 (a-d) displays screenshots of stages of BoTySeGa. The first three stages comprise of an opening stage, a comment stage revealed in case of player sent a wrong solution, and a hint stage to assist player who access Help feature. The fourth stage displays gameplay data window.

C. Data Type and Value of Gameplay

Score is not the only attribute recorded from BoTySeGa players. BoTySeGa also records completion time of a challenge and the frequency of accessing Help. We call the records as gameplay data representing players' behavior in playing BoTySeGa.

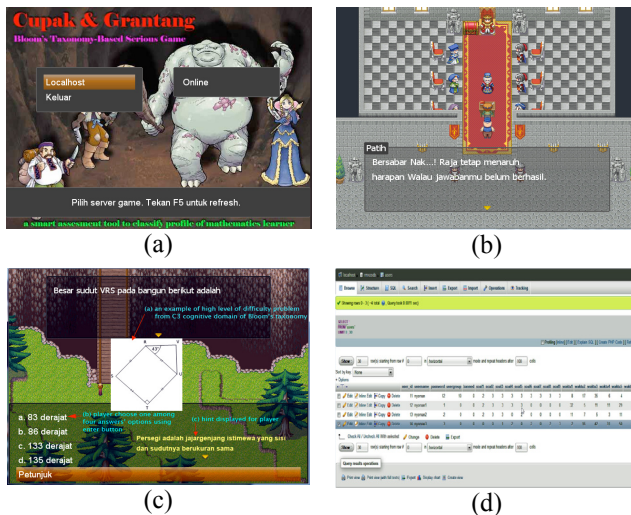


Figure 3. Screenshot Stages of BoTySeGa

A player score is defined as a measure of player's mastery on learning content as the basis of determining player's progress or his competence. Player score is recorded in ordinal data type 0, 1, 2, 3. A value 0 represents inability of the player to solve the problem correctly. Value 1, 2, 3 successively represent the success of the player in solving any problem in any level of difficulty.

Completion time represents duration since appearance of a challenge in the serious game up to the player's submission of a solution. Data are recorded in ratio data type. Measurement unit of completion time is in second (sec). Theoretically, serious game did not constrain completion time spent for a challenge. Learner's understanding is defined as learner's experience in achieving learning objective. Timing the game completion constrain the players' performance since they are forced to decide a solution a hurry and risky [4]. The longer the time completion, the more accurate will the solution be performed by the player [20]. Conversely, [21, 22] state that response time of test item is influenced by intelligence that intellectual ability positively correlates with person speed in processing information. Referring to [4, 20, 21, 22], we restrict the time completion time with a value of 10 minutes (600 sec). It representing a very long completion time and it is never revealed as BoTySeGa gameplay data.

The number of accessing Help represents frequency of the player in accessing Help feature. BoTySeGa records the time when the player accesses Help feature. Data are recorded in ratio data type and unit of measurement is sec. The value n represents the player accesses Help at the n^{th} sec.

Data of players' achievement are recorded in ratio data type using paper-based test. Value of the player's achievement ranges from 0 to 10. Value 0 means the player's low achievement, while 10 represents the player's high achievement.

IV. RESEARCH METHODOLOGY

Research methodology described in this section includes gameplay data collection, pre-process, and classification of gameplay data as shown in Fig. 4.

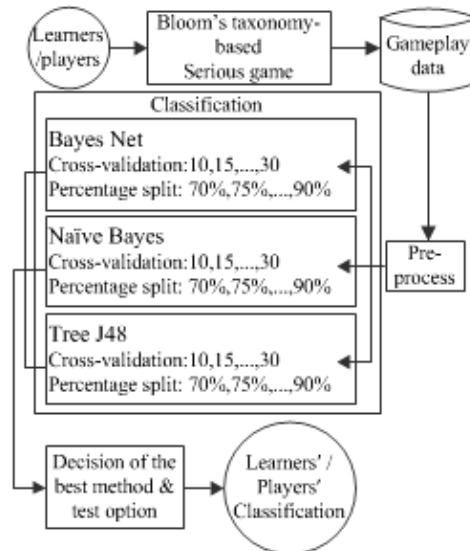


Figure 4. Research Methodology

A. Gameplay Data Collection

The players' data collection was done at computer laboratory of Universitas Pendidikan Ganesha. As many as 85 5th grade elementary students participated as research subjects. Data were collected through the operation of ten stand-alone personal computers. We divided research subjects into three groups. Each group consisted of 17-20 research subjects. Data of each group were taken in one day for about 6 hours. The maximum time allotted for 10 players or less was two hours. Data were collected from 8.00 a.m. until 02.00 p.m. Installation of ten personal computers for data collection was purposefully intended to guarantee accuracy of data recorded showing serious game players' behavior. Two research observers were involved during the collection of research data.

In the beginning; the players were informed that their behavior in playing serious game would be recorded as gameplay data. They were also informed that they themselves were in charge of playing character's tasks in winning the game. Furthermore; they were informed about the manual operation of BoTySeGa and asked to perform login process before playing the game. The two research personals present during data collection were assigned to supervise the players to maximize the use of existing features, focus on their main task, or allocate as short as possible time to complete a challenge. This step was held to supervise players that they could concentrate on their tasks to win the game. It also limits the wasted time during a game play which tends to be performed by procrastinator. Players who had completed all challenges or resigned before completing the whole challenges were obliged to communicate their decision by clicking Quit feature available in the serious game.

Data on the players' achievement were taken based on paper-based test. The data were collected on the similar day after the players playing BoTySeGa. We also collected data from experts; in our study they were teachers of the learners. Data from experts were used as reference data or ground truth data of the players' achievement. Divergence of the achievement values (if exist) taken from the paper-based test and the teachers were discussed among the researchers, the teachers and the learners.

B. Pre-process of Gameplay Data

There are three aims of pre-process stage. First, pre-process closely deals with players' data which do not completely solve nine challenges of BoTySeGa. The second and third phase focus on determining how often the players access Help and creating the actual knowledge based on score data.

In case of game over state, attribute completion time will record a default value 0. Surely, it does not give a representative result. The player who does not submit solution is labeled spent a shortest time to solve a challenge. For the case, we set a value 600 of which is never revealed as empirical data of the players' completion time.

Dealing with Help attribute; this paper focuses on how frequent activities of accessing of Help is done. It means that we need to convert the players' data into a value representing how often Help feature is accessed before the players submit a solution in the game.

The following pre-process we did was determining the players' actual knowledge value. This value was based on players' score for each challenge. The maximum value of the players' actual knowledge is 9. A player who failed in solving challenges in a cognitive domain was subject to be considered having no actual knowledge in that domain. We reduced the actual knowledge value of the players by one every time they failed in solving a challenge. Procedure to determining the players' actual knowledge was done as in algorithm shown below to create actual knowledge value.

```

Algorithm to create actual knowledge value
Set n to 0
For all challenges
  If not (solution is right) then
    Increase (n)
Set AK to 9 - n

```

An example of gameplay data before versus after pre-process is shown in Table III.

The last aim of pre-process is classifying class target. We selected players' achievement data as class target and classified it into three classes: Poor, Fair, and Good. Twenty seven percent of lower group is classified as Poor class, Twenty seven percent of higher group is classified as Good class, and others are classified as Fair class.

C. Classification of Gameplay Data

Process of classification of gameplay data is given in the following.

- Classify gameplay data by applying two tests option namely: cross-validation and percentage split. Classifications are done in a variety of numbers of folds: 10, 15, 20, 25, and 30 and percentages of split: 70%, 75%, 80%, 85%, and 90%.
- Check the classification results that meet the strength of agreement category which at least is Good ($0.61 \leq \text{Kappa value } (\kappa) \leq 0.80$).
- Determine number of fold or percentage of split that give maximum percentage of correctly classified instances for the three methods.
- Determine optimum percentage of correctly classified instances for the fold or percentage of split given from the previous point.

TABLE III.
BEFORE VS. AFTER PRE-PROCESS OF BOTYSEGA PLAYERS' DATA

Attributes	Before	After			Description
		First	second	Third	
S ₂	0	0	0	0	Case on Player's data who stated game over after the sixth challenge.
S ₃	2	2	2	2	
S ₄	0	0	0	0	
S ₅	2	2	2	2	
S ₆	0	0	0	0	
S ₇	0	0	0	0	
S ₈	0	0	0	0	
S ₉	0	0	0	0	
t ₁	36	36	36	36	
t ₂	30	30	30	30	
t ₃	40	40	40	40	
t ₄	45	45	45	45	
t ₅	50	50	50	50	
t ₆	65	65	65	65	
t ₇	0	600 ^{*)}	600	600	
t ₈	0	600 ^{*)}	600	600	
t ₉	0	600 ^{*)}	600	600	
A2H ₁	0	0	0	0	
A2H ₂	0	0	0	0	
A2H ₃	0	0	0	0	
A2H ₄	40	40	1 ^{*)}	1	
A2H ₅	30,45	30,45	2 ^{*)}	2	
A2H ₆	50	50	1 ^{*)}	1	
A2H ₇	0	0	0	0	
A2H ₈	0	0	0	0	
A2H ₉	0	0	0	0	
AK _n				3 ^{*)}	
Ach	3	3	3	3	

^{*)} modified value during pre-process

V. RESULTS AND DISCUSSION

In this section we show and discuss the performance of three different classification methods described above. Each method was implemented for a similar gameplay data from 85 instances. In this paper, an instance represents the gameplay data of one player. Each of the instances involved 29 attributes. Before the implementation of each method, we conducted pre-process for gameplay data following the procedure described in the research methodology section. The following points describe classification results of each method based on two test options, and comparison of the results at an optimum cross-validation or percentage split test option.

A. The Classification Results

The results of classification implementation of the above three classification methods and two tests option is shown in Table IV. The Performance evaluation of classification was measured by the accuracy of classification method which were categorized at least within the level of Good strength of agreement category.

Referring to Table IV, we found three results classified at Very Good strength of agreement category ($0.81 \leq \kappa \leq 1.00$). One of the results was obtained from the classification implementation of NB with 85% of split test option. The accuracy of the result is 92.31%. The other two results were obtained from J48 with 80% and 85% percentages of split test option. The accuracy of those classifications is 94.12% and 92.31%. Five results are classified as Good ($0.61 \leq \kappa \leq 0.80$). Two were obtained from the classification implementation of BN with 85% and 90% of split test. The accuracy of those classifications is successively 84.62% and 87.65%. The other three re-

sults were from NB with 75%, 80%, and 90% percentages of split test. The accuracy of the last three results successively is: 85.71%, 88.24%, and 87.50%. We also found that the percentage of split 85% is the best percentage for the three classification methods.

TABLE IV.
PERCENTAGE OF CORRECTLY CLASSIFIED INSTANCES, KAPPA VALUES,
AND MEAN ABSOLUTE ERRORS OF CLASSIFICATION RESULTS

		BN			NB			J48		
		CC	κ	MAE	CC	κ	MAE	CC	κ	MAE
Fold	10	63.53	0.41	0.24	68.24	0.47	0.21	63.53	0.35	0.28
	15	65.88	0.45	0.22	69.41	0.48	0.21	64.71	0.37	0.27
	20	64.71	0.43	0.23	69.41	0.49	0.21	61.18	0.30	0.28
	25	64.71	0.44	0.23	68.24	0.47	0.22	63.53	0.33	0.27
	30	67.06	0.47	0.22	67.06	0.45	0.22	60.00	0.28	0.30
Per- cent (%)	70	80.00	0.64	0.14	80.00	0.64	0.14	60.00	0.29	0.27
	75	76.19	0.59	0.18	85.71	0.74	0.09	76.19	0.51	0.17
	80	76.47	0.59	0.17	88.24	0.78	0.07	94.12	0.88	0.16
	85	84.62	0.72	0.11	92.31	0.85	0.05	92.31	0.84	0.15
	90	87.50	0.71	0.13	87.50	0.71	0.08	75.00	0.47	0.25

B. Comparison Results of the Classification methods at Percentage of Split 85%.

Comparison of correctly classified instances, Kappa value, and mean absolute error for the three classification methods is depicted in Fig. 5.

Based on the classification results, both NB and J48 classifiers give the similar percentage of correctly classified instances (92.31%). In view of Kappa value, both classification results are categorized as Very Good NB, however, has less MAE than J48.

Visualization of classification error for NB classifier with 85% of percentage split is depicted in Fig. 6. One out of thirteen instances of gameplay data was labeled incorrectly class. It was labeled Fair for the instance that should be predicted Poor.

The result depicted in Fig. 5 shows that NB outperforms BN and J48 classifiers with respect to the accuracy and mean absolute error of classification of BoTySeGa gameplay data. Our result confirms the previous result [23, 24] that states NB classifier should be favored over BN and J48. Detail of classification results of fifteen percent of testing data is shown in Table V.

VI. CONCLUSION

We applied three different classifiers BN, NB and J48 to classify BoTySeGa gameplay data. Two test options

were chosen in each of the classification namely: cross-validation and percentage split. We chose 10,15,20,25, and 30 fold of cross-validation for test option and 70, 75, 80, 85, and 90 percent of percentage split. These classifications aimed at finding the best classifier and procedure of classification that provides high accuracy value for the classification of data BoTySeGa gameplay data. Result shows that NB classifier with 85% split test option outperforms other two classifiers BN and J48. The accuracy value of this classifier is 92.31% and its MAE is 0.05. This implies that the usage of BoTySeGa as an assessment tool replacing paper-based should be considered as a new pedagogical issue.

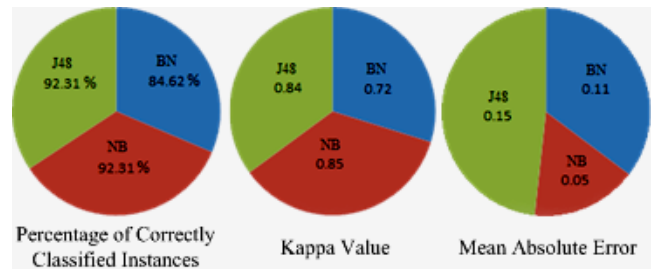


Figure 5. Percentage of correctly classified instances, Kappa value, and MAE of BN, NB, and J48 for the percentage of split 85%

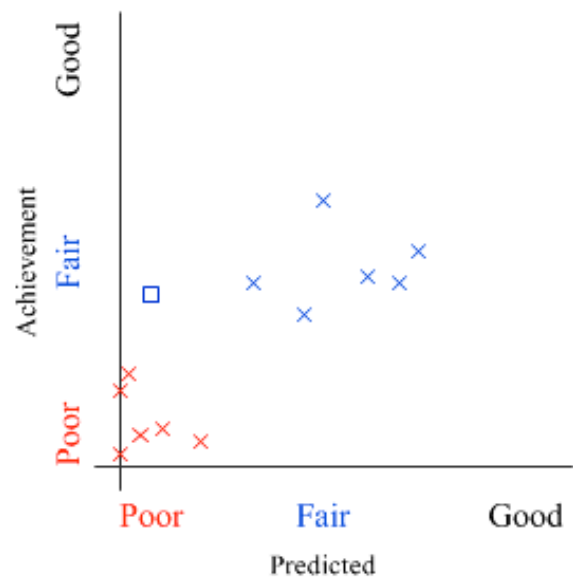


Figure 6. Visualization of Classification Error for NB Classifier with 85% Percentage of Split

TABLE V CLASSIFICATION RESULTS OF 15% TESTING DATA

No	s ₁	s ₂	s ₃	s ₄	s ₅	s ₆	s ₇	s ₈	s ₉	t ₁	t ₂	t ₃	t ₄	t ₅	t ₆	t ₇	t ₈	t ₉	A2H									AK n	Pred. Ach	ACh		
																			1	2	3	4	5	6	7	8	9					
1	0	0	1	2	0	2	3	0	0	30	9	21	9	35	192	103	108	22	0	0	0	0	0	0	1	1	1	1	4	Fair	Fair	
2	0	1	2	0	2	3	3	0	0	15	3	40	85	10	23	12	28	600	0	0	0	0	0	0	0	0	0	0	5	Fair	Fair	
3	0	0	1	0	0	1	0	1	2	28	39	108	143	11	14	65	100	125	0	1	0	0	0	0	0	0	0	1	1	4	Fair	Fair
4	0	1	2	0	0	1	2	3	3	18	38	25	33	9	15	62	230	98	0	1	0	0	0	0	0	1	0	0	6	Fair	Fair	
5	0	0	1	2	0	0	0	0	0	25	27	5	40	105	600	600	600	600	0	0	0	0	0	0	0	0	0	0	2	Poor	Poor	
6	0	1	2	0	0	0	0	0	0	35	75	65	85	100	135	600	600	600	0	0	0	0	1	0	1	0	0	0	2	Poor	Poor	
7	0	0	1	0	0	1	0	0	1	159	49	51	8	12	332	150	36	32	0	0	0	0	0	0	0	0	0	0	3	Fair	Fair	
8	2	0	2	0	0	1	0	0	0	7	21	39	11	5	6	23	600	600	0	0	0	0	0	0	0	0	0	0	3	Poor ^{*)}	Fair	
9	2	3	0	0	0	0	0	0	0	83	54	70	600	600	600	600	600	600	1	0	1	0	0	0	0	0	0	0	2	Poor	Poor	
10	2	3	3	3	0	2	0	0	0	42	8	34	91	139	82	162	142	600	0	0	0	0	0	0	0	0	0	0	5	Fair	Fair	
11	2	0	2	0	2	0	0	0	0	19	15	59	140	11	9	600	600	600	0	0	1	1	0	0	0	0	0	0	3	Poor	Poor	
12	0	0	1	0	0	0	0	0	0	18	44	47	29	600	600	600	600	600	0	0	0	0	0	0	0	0	0	0	1	Poor	Poor	
13	0	0	0	0	0	0	0	0	0	10	139	32	600	600	600	600	600	600	0	0	0	0	0	0	0	0	0	0	0	Poor	Poor	

ACKNOWLEDGMENT

The authors convey greatest gratitude to the Directorate General of Higher Education which provided scholarship for our study and Sandwich-like program at Institute for Mathematics and Computing Science, University of Groningen, Netherlands. We also thank to 5th grade of elementary school students and teachers of SD 3 Banjar Jawa and SD Laboratorium Singaraja who had participated in gameplay data collection.

REFERENCES

- [1] M. Nakayama and R. Santiago, "Learner Characteristics and Online Learning," in *Encyclopedia of the Sciences of Learning*, Springer, 2012, pp. 1745–1747.
- [2] M. Samarakou, A. Papadakis, E. D. Fylladitakis, A. HatziaPOSTOLOU, G. Tsaganou, and W. G. Früh, "An Open Learning Environment for the Diagnosis, Assistance and Evaluation of Students Based on Artificial Intelligence," *Int. J. Emerg. Technol. Learn. IJET*, vol. 9, no. 3, pp. 36–44, May 2014.
- [3] F. Bellotti, B. Kapralos, K. Lee, and P. Moreno-Ger, "User Assessment in Serious Games and Technology-Enhanced Learning," *Adv. Hum.-Comput. Interact.*, vol. 2013, p. e120791, Mar. 2013. <http://dx.doi.org/10.1155/2013/120791>
- [4] C. S. Loh and Y. Sheng, "Maximum Similarity Index (MSI): A metric to differentiate the performance of novices vs. multiple-experts in serious games," *Comput. Hum. Behav.*, vol. 39, pp. 322–330, Oct. 2014. <http://dx.doi.org/10.1016/j.chb.2014.07.022>
- [5] "Proof of Learning: Assessment in Serious Games." [Online]. Available: http://www.gamasutra.com/view/feature/2433/prof_of_learning_assessment_in_php. [Accessed: 30-Sep-2014].
- [6] R. Van Eck, "Digital game-based learning: It's not just the digital natives who are restless," *Educ. Rev.*, vol. 41, no. 2, p. 16, 2006.
- [7] J. Kirkley, S. Kirkley, and J. Heneghan, "Building bridges between serious game design and instructional design," *Des. Use Simul. Comput. Games Educ.*, pp. 61–83, 2007.
- [8] G. Wallner, "Play-Graph: A methodology and visualization approach for the analysis of gameplay data," in *FDG*, 2013, pp. 253–260.
- [9] C. Conati, "Building and Evaluating an Intelligent Pedagogical Agent to Improve the Effectiveness of an Educational Game," *PROC 9TH INT CONF INTELL. USER INTERFACE ACM PRESS*, pp. 6–13, 2004.
- [10] M. A. Syufagi, M. Hariadi, and M. H. Purnomo, "A Cognitive Skill Classification Based on Multi Objective Optimization Using Learning Vector Quantization for Serious Games," *ITB J. Inf. Commun. Technol.*, vol. 5, no. 3, pp. 189–206, 2011.
- [11] N. Yusof, N. Bahiah, M. Shahizan, and Y. Chun, "A Concise Fuzzy Rule Base to Reason Student Performance Based on Rough-Fuzzy Approach," in *Fuzzy Inference System - Theory and Applications*, M. F. Azeem, Ed. InTech, 2012.
- [12] N. E. Dunbar, C. H. Miller, B. J. Adame, J. Elizondo, S. N. Wilson, B. L. Lane, A. A. Kauffman, E. Bessarabova, M. L. Jensen, S. K. Straub, Y.-H. Lee, J. K. Burgoon, J. J. Valacich, J. Jenkins, and J. Zhang, "Implicit and explicit training in the mitigation of cognitive bias through the use of a serious game," *Comput. Hum. Behav.*, vol. 37, pp. 307–318, Aug. 2014. <http://dx.doi.org/10.1016/j.chb.2014.04.053>
- [13] J. Feldman, A. Monteserin, and A. Amandi, "Detecting students' perception style by using games," *Comput. Educ.*, vol. 71, pp. 14–22, Feb. 2014. <http://dx.doi.org/10.1016/j.compedu.2013.09.007>
- [14] M. Minović, M. Milovanović, U. Šošević, and M. Á. Conde González, "Visualisation of student learning model in serious games," *Comput. Hum. Behav.*
- [15] T. Highley and A. E. Edlin, "Discrete Mathematics assessment using learning objectives based on Bloom's taxonomy," in *39th IEEE Frontiers in Education Conference, 2009. FIE '09*, 2009, pp. 1–6.
- [16] C. S. Tan and Y. H. Tay, "A preliminary study of the progress on using Bayesian Networks for educational assessment," in *2010 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT)*, 2010, pp. 75–80.
- [17] B. S. Bloom, *Taxonomy of educational objectives: the classification of educational goals*. D. McKay, 1956.
- [18] H. Chauhan, V. Kumar, S. Pundir, and E. S. Pilli, "A Comparative Study of Classification Techniques for Intrusion Detection," in *2013 International Symposium on Computational and Business Intelligence (ISCBI)*, 2013, pp. 40–43.
- [19] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition*, 3 edition. Morgan Kaufmann, 2011.
- [20] G. Hawkins, K. Nesbitt, and S. Brown, "Dynamic Difficulty Balancing for Cautious Players and Risk Takers," *Int. J. Comput. Games Technol.*, vol. 2012, p. e625476, Jun. 2012. <http://dx.doi.org/10.1155/2012/625476>
- [21] A. R. Jensen, *Bias in Mental Testing*. New York: Free Press, 1980.
- [22] A. R. Jensen, "Reaction Time and Psychometric g," in *A Model for Intelligence*, P. H. J. Eysenck, Ed. Springer Berlin Heidelberg, 1982, pp. 93–132.
- [23] A. Cufoglu, M. Lohi, and K. Madani, "Classification accuracy performance of Naïve Bayesian (NB), Bayesian Networks (BN), Lazy Learning of Bayesian Rules (LBR) and Instance-Based Learner (IB1) - comparative study," 2008, pp. 210–215.
- [24] A. Cufoglu, M. Lohi, and K. Madani, "A Comparative Study of Selected Classifiers with Classification Accuracy in User Profiling," presented at the Computer Science and Information Engineering, 2009 WRI World Congress on, 2009, vol. 3, pp. 708–712.

AUTHORS

I Nyoman Sukajaya is with Mathematics Department of Education, Faculty of Mathematics and Natural Sciences, Universitas Pendidikan Ganesha, Singaraja-Indonesia (e-mail: nyoman11@mhs.ee.its.ac.id). He received Bachelor degree in Mathematics of Education at Udayana University in 1991 and master degree in Informatics Engineering at Faculty of Industrial engineering, Institut Teknologi Bandung, Bandung in 1999. Since 2011, he is pursuing a Ph.D degree at electrical engineering Department, Institut Teknologi Sepuluh Nopember (ITS), Surabaya-Indonesia. His research topic of interests is the implementation of information technology in learning.

I Ketut Eddy Purnama was with Electrical Engineering Department, (e-mail: ketut@ee.its.ac.id) and now he is head and Associate Professor of Multimedia and Network Engineering Department, Institut Teknologi Sepuluh Nopember (ITS), Surabaya-Indonesia (e-mail: ketut@te.its.ac.id). He received bachelor degree in Electrical Engineering at ITS in 1994. His Master degree is in Informatics Engineering, Faculty of Industrial Technology, Institut Teknologi Bandung, Bandung, in 1999, and Ph.D degree from University of Groningen, the Netherlands in 2007. His research interest is in Data Mining, Medical Image Processing and Intelligent System.

Mauridhi Hery Purnomo is with the Electrical Engineering Department, Institut Teknologi Sepuluh Nopember (ITS), Surabaya-Indonesia (e-mail: hery@ee.its.ac.id). He received bachelor degree from ITS, in 1985. His M.Eng., and Ph.D degrees was received from Osaka City University, Osaka, Japan in 1995, and 1997, respectively. He joined ITS in 1985 and has been a Professor since 2003. His current interests include intelligent system applications, image processing, medical imaging, control and management. He is a Member of IEEE and INNS.

Submitted 04 February 2015. Published as resubmitted by the authors 24 March 2015.